



**Europäisches
Patentamt**

**European
Patent Office**

**Office européen
des brevets**

Bescheinigung

Certificate

Attestation

Die angehefteten Unterla-
gen stimmen mit der
ursprünglich eingereichten
Fassung der auf dem näch-
sten Blatt bezeichneten
europäischen Patentanmel-
dung überein.

The attached documents
are exact copies of the
European patent application
described on the following
page, as originally filed.

Les documents fixés à
cette attestation sont
conformes à la version
initialement déposée de
la demande de brevet
européen spécifiée à la
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

01402546.4

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

R C van Dijk

THIS PAGE BLANK (USPTO)



Anmeldung Nr:
Application no.: 01402546.4
Demande no:

Anmeldetag:
Date of filing: 02.10.01
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

Koninklijke Philips Electronics N.V.
Groenewoudseweg 1
5621 BA Eindhoven
PAYS-BAS

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.
If no title is shown please refer to the description.
Si aucun titre n'est indiqué se referer à la description.)

Video decoding method and corresponding decoder

In Anspruch genommene Priorität(en) / Priority(ies) claimed /Priorité(s)
revendiquée(s)

Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/
Classification internationale des brevets:

H04N7/26

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of
filing/Etats contractants désignées lors du dépôt:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE TR

THIS PAGE BLANK (USPTO)

«VIDEO DECODING METHOD AND CORRESPONDING DECODER»

FIELD OF THE INVENTION

5 The present invention generally relates to video decompression, and more particularly to a decoding method for decoding a video bitstream including base layer coded video signals and enhancement layer coded video signals and generating decoded signals corresponding either only to the base layer signals, to be then displayed alone, or to the base layer signals and the enhancement layer signals, to be then displayed
10 together. It also relates to a corresponding video decoder.

BACKGROUND OF THE INVENTION

 The MPEG-4 video standard is described for instance in the document "Overview of the MPEG-4 Version 1 Standard", ISO/IEC JTC1/SC29/WG11 N1909, October 1997, Fribourg. In an MPEG-4 encoder, three types of pictures are used : intra-coded (I) pictures, coded independently from other pictures, predictively-coded (P) pictures,
15 predicted from a past reference picture (I or P) by motion compensated prediction, and bidirectionally predictively-coded (B) pictures, predicted from a past and a future reference picture (I or P). The I pictures are the most important, since they are reference pictures ; moreover, they can provide access points (in the bitstream) where decoding
20 can begin without any reference to previous pictures. In such pictures, only the spatial redundancy is eliminated. Usually they are the VOPs that are most expensive in terms of bits (a VOP is a Video Object Plane, i.e. an instance of Video Object at a given time, a VO being itself a specific object inside a scene).

 By reducing both spatial and temporal redundancy, P-pictures offer a better
25 compression compared to I-pictures which reduce only the spatial redundancy. B-pictures offer the highest degree of compression, and a B-VOP therefore cannot be a reference VOP. As for a P-VOP, only the difference between the current B-VOP and its reference VOP(s) is coded. Only P- and B-VOPs are concerned by the motion estimation, carried out according to the so-called "Block Matching Algorithm" : for each macroblock of the
30 current frame, the macroblock which matches the best in the reference VOP is sought in a predetermined search zone, and a motion vector MV is then calculated. The resemblance criterion is given by the Sum of Absolute Differences (SAD). For a NxN macroblock, SAD is expressed as :

$$SAD = \sum_{i=0}^{NxN} |A(i) - B(i)|$$

Thus the chosen macroblock is the one corresponding to the smallest SAD among those calculated in the search zone. For said estimation, different modes exist, depending on the type of the VOP :

- (a) for P-VOPs macroblocks, only the "forward mode" (use of a past reference I-VOP or P-VOP) is available ;
- (b) for B-VOPs macroblocks, four modes are available for the macroblock estimation :
 - "forward mode" (as for P-VOPs) ;
 - "backward mode" : as the forward mode, except that the reference is no longer a past one but a future P- or I-VOP ;
 - "interpolated mode" or "bidirectional mode" : it combines the forward and backward modes and uses a past and a future reference VOP ;
 - "direct mode" : each motion vector is calculated thanks to the motion vector of the future reference VOP and thanks to the temporal distance between the different VOPs.

Within MPEG-4, an important functionality, the scalability, is offered. Scalable coding, also known as "layered coding", allows to generate a coded representation in a manner that enables a scalable decoding operation. Scalability is the property of a bitstream to allow decoding of appropriate subsets of data leading to the generation of complete pictures of resolution and/or quality that commensurate with the proportion of the bitstream decoded. Such a functionality is useful in the numerous applications that require video sequences to be simultaneously available at a variety of resolutions and/or quality and/or complexity. Indeed, if a bitstream is scalable, one user will access only a portion of it to provide basic video in accordance with his own decoder or display, or with the available bandwidth, while another one will use the full bitstream to produce a better video quality.

The advantage of scalability, which costs less in terms of coding process than the solution according to which several independent bitstreams are coded, is that it allows to deliver a bitstream separable into at least two different bitstreams (and, among them, one with a higher bitrate than the others). Each type of scalability therefore involves more than one layer. In the case of temporal scalability, at least two layers consisting of a lower layer and a higher layer are considered. The lower layer is referred to as the base layer, encoded at a given frame rate, and the additional layer is called the enhancement layer, encoded to provide a higher temporal resolution at the display side. A decoder may decode only the base layer, which corresponds to the minimum amount of data required to decode the video stream, or also decode the enhancement layer (in addition to the base layer), said enhancement layer corresponding to the additional data required to provide, if associated to the data corresponding to the base layer, an

enhanced video signal, and then output more frames per second if a higher resolution is required.

However, at the decoding side, there are situations where a large difference of quality between the displayed images of the base layer and those of the enhancement layer is observed, for example when the available bandwidth for each layer is very different. In that case, the subjective quality of the decoded sequence can be quite low because of the flickering effect, even if only a few frames (those of the base layer) have a significantly lower quality, compared with the average of the sequence.

SUMMARY OF THE INVENTION

It is therefore the object of the invention to propose a video decoding method allowing to improve the quality of the displayed decoded sequence.

To this end the invention relates to a decoding method such as defined in the introductory paragraph of the description and comprising the steps of :

- decoding the base layer coded video signals to produce decoded base layer frames ;
 - decoding the enhancement layer coded video signals to produce decoded enhancement layer frames ;
 - displaying the decoded base layer frames either alone or with the decoded enhancement layer frames to form video frames ;
- said method being characterized in that the displaying step itself comprises :
- a decision sub-step, for examining on the basis of a given criterion the quality of each successive base layer frame to be displayed and selecting the poor quality frames ;
 - a replacement sub-step, for replacing each poor quality base layer frame by at least one of the two frames of the enhancement layer preceding and following said poor quality frame base layer frame.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will now be described in a more detailed manner, with reference to the accompanying drawing in which Fig.1 shows the general implementation of a system for coding and decoding a video sequence.

DETAILED DESCRIPTION OF THE INVENTION

A system for coding and decoding a video sequence is generally implemented as shown in Fig.1. Said system comprises a video encoding part 1, a video decoding part 3 and, between them, a transmitting medium 2. The encoding part 1 comprises a video frame source 11 generating uncompressed video frames, a video encoder 12 provided for coding the frames it receives from the source 11, and an encoder buffer 13. In the encoder 12, the uncompressed video frames entering at a given frame

rate are coded according to the principles of the MPEG-4 standard and transmitted to the encoder buffer 13 at the output of which the stored, coded frames are sent towards the transmitting medium 2.

At the decoding side, the transmitted coded frames are received by the video decoding part 3 which comprises a decoder buffer 14, a video decoder 15 and a video display 16. The decoder buffer 14 receives and stores the transmitted, coded frames and itself transmits them to the video decoder 15 which decodes these frames, generally at the same frame rate. The decoded frames are then sent to the video display 16 which displays them.

In the present case of a scalable coding scheme, the video encoder 12 comprises a base layer encoding stage, which receives from the source 11 the frames corresponding to the original video signal and codes the frames for generating a base layer bitstream sent to the encoder buffer 13, and an enhancement layer encoding stage, which receives on the one hand (from the source 11) the frames corresponding to the original video signal and on the other hand decoded frames derived from the coded frames transmitted in the base layer bitstream. This enhancement layer encoding stage generates, in the form of an enhancement layer coded bitstream, a residual signal that represents the image information missing in the base layer frames and may therefore be added to the base layer bitstream.

Reciprocally, on the decoding side, the decoder 15 of the video decoding part 3 comprises processing circuitry provided for receiving the coded base layer bitstream and the coded enhancement layer bitstream and sending towards the video display 16 decoded signals corresponding either to the base layer signals, then displayed alone, or to the base layer signals associated with the enhancement layer signals, displayed together.

Under some conditions, and for instance when the available bandwidth for each layer is very different, a very large difference of quality between the displayed images coming from the base layer and the displayed images coming from the enhancement layer is observed. In such a situation, the subjective quality of the displayed, decoded sequence will be low, owing to a flickering effect, even if only a few frames in the base layer have a significantly lower quality compared with the average quality of the sequence. This drawback may be avoided if said poor quality frames of the base layer are not displayed and are replaced by frames coming from the enhancement layer.

These replacement frames may be for example frames interpolated from the preceding and following frames of the enhancement layer. The replacement frame may also be obtained by copying one of said preceding and following frames, for instance the temporally closest one.

For deciding whether the decoded frames have a sufficient quality to be displayed, a quantitative criterion has to be defined. It is for instance possible to store and compare the quantization step sizes of the successive frames : in case of a very noticeable difference of said step size for a frame with respect to the other preceding and following frames, it is likely that said frame has a poor quality. Another criterion may be the following. Each frame being divided into 8×8 blocks, the texture gradient at the boundaries of said blocks is examined : if said gradient is noticeably higher in a specific base layer frame, said frame is considered as having a poor quality and is not displayed.

It must be understood that the video decoder described herein can be implemented in hardware or software, or by means of a combination of hardware and software. It may then be implemented by any type of computer system or other apparatus adapted for carrying out the method described hereinabove, comprising for instance a memory which stores computer-executable process steps and a processor which executes the process steps stored in the memory so as to produce the decoded frames to be displayed. A typical combination of hardware and software could be a general-purpose computer system with a computer program that, when loaded and executed, controls the computer system such that it carries out the method described herein. Alternatively, a specific use computer, containing specialized hardware for carrying out one or more of the functional tasks of the invention, could be utilized. The present invention can also be embedded in a computer program medium or product, which comprises all the features enabling the implementation of the method and functions described herein, and which - when loaded in a computer system - is able to carry out these method and functions. The invention also relates to the computer executable process steps stored on such a computer readable medium or product and provided for carrying out the described video decoding method. Computer program, software program, program, program product, or software, in the present context mean any expression, in any language, code or notation, of a set of instructions intended to cause a system having an information processing capability to perform a particular function either directly or after either or both of the following: (a) conversion to another language, code or notation, and/or (b) reproduction in a different material form.

The foregoing description of the invention has been presented for purposes of illustration and description, and is not intended to be exhaustive or to limit the invention to the precise form disclosed, and modifications and variations are possible in light of the above teachings. Such modifications and variations that are apparent to a person skilled in the art are intended to be included within the scope of this invention as defined by the accompanying claims.

THIS PAGE BLANK (USPTO)

CLAIMS :

1. For use in a video decoder comprising processing circuitry capable of receiving from a transmitting and/or storing medium a video bitstream which itself includes base layer coded video signals and enhancement layer coded video signals and decoding said bitstream for generating decoded signals corresponding either only to the base layer signals, to be then displayed alone, or to the base layer signals and the enhancement layer signals, to be then displayed together, a method of decoding said video bitstream including said base layer and enhancement layer coded video signals, comprising the steps of :
 - decoding the base layer coded video signals to produce decoded base layer frames ;
 - decoding the enhancement layer coded video signals to produce decoded enhancement layer frames ;
 - displaying the decoded base layer frames either alone or with the decoded enhancement layer frames to form video frames ;
 said method being characterized in that the displaying step itself comprises :
 - a decision sub-step, for examining on the basis of a given criterion the quality of each successive base layer frame to be displayed and selecting the poor quality frames ;
 - a replacement sub-step, for replacing each poor quality base layer frame by at least one of the two frames of the enhancement layer preceding and following said poor quality frame base layer frame.
2. A decoding method according to claim 1, in which each poor quality base layer frame is replaced by the temporally closest of said preceding and following frames of the enhancement layer.
3. A decoding method according to claim 1, in which said poor quality base layer frame is replaced by a frame obtained by means of an interpolation between said preceding and following frames of the enhancement layer.
4. A video decoder for decoding a video bitstream including base layer coded video signals and enhancement layer coded video signals, wherein the enhancement layer includes enhancement frames arranged in a display order, said decoder comprising:
 - first decoding means for producing decoded base layer frames ;
 - second decoding means for producing decoded enhancement layer frames ;
 - displaying means for displaying said decoded base layer and enhancement layer frames and in which each poor quality frame of the base layer to be displayed is replaced by an frame obtained either by means of an interpolation between the two frames of the enhancement layer preceding and following said poor quality frame of the base layer or by only one of these two frames.

THIS PAGE BLANK (USPTO)

Abstract

The invention relates to a method of decoding a video bitstream including base layer and enhancement layer coded video signals, said method comprising the steps of decoding the base layer and enhancement layer coded video signals to produce decoded base layer frames and decoded enhancement layer frames, and displaying the decoded base layer frames either alone or with the decoded enhancement layer frames. According to the invention, each poor quality frame of the base layer to be displayed is replaced by an frame obtained either by means of an interpolation between the two frames of the enhancement layer preceding and following said poor quality frame of the base layer or by only one of these two frames, for example the temporally closest one.

THIS PAGE BLANK (USPTO,

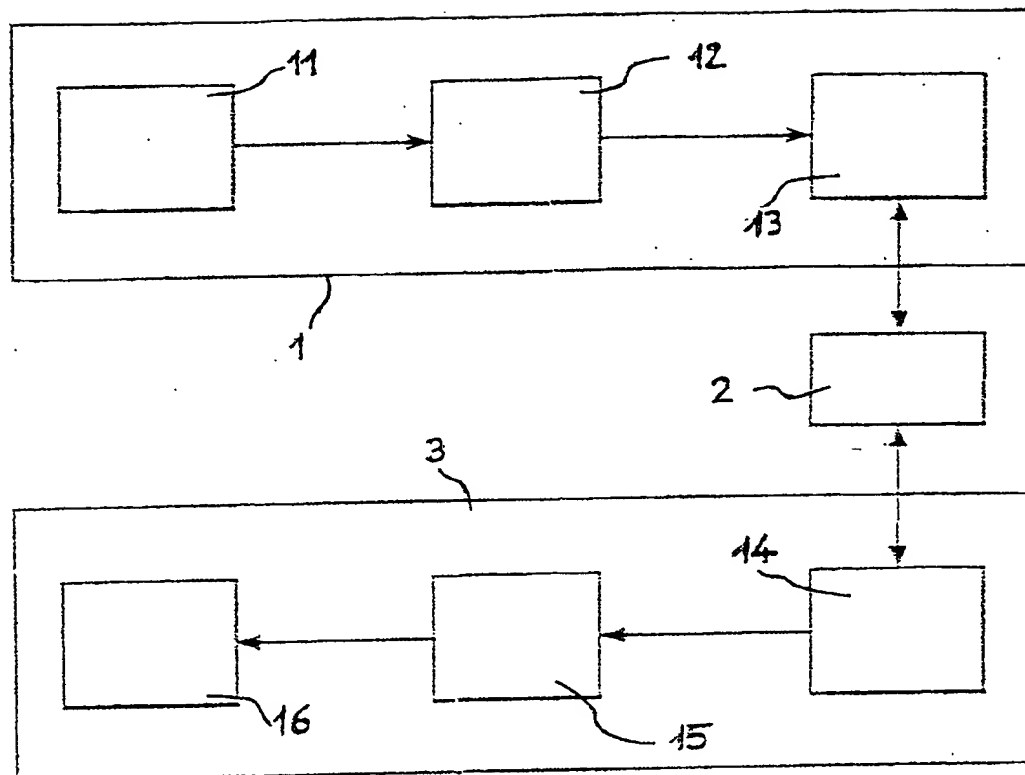


FIG. 1

THIS PAGE BLANK (USPTO)